METHOD FOR DETECTING A PROPENSITY OF AN INDIVIDUAL TO RESPONSE EFFECTIVELY TO TREATMENT OF INTERFERON-α AND RIBAVIRIN COMBINED THERAPY

FIELD OF THE INVENTION

The present invention relates generally to HCV treatment, and more particularly to gene and genetic polymorphisms associated with treatment efficacy of interferon-α and ribavirin combined therapy on HCV patients.

BACKGROUND OF THE INVENTION

According to most recent World Health Organization (WHO) estimates, around 170-200 million individuals have chronic HCV infection worldwide. The HCV prevalence shows significant geographic variations and demographic variations within a geographic area. In Europe, and particularly in the Mediterranean countries, the prevalence of HCV infection increases in parallel with age, while in the United States it is most common in persons 30-49 years of age, see W.R. Kim,

The burden of hepatitis C in the United States, Hepatology 36 (2002) S30-S34. The incidence of new infections with HCV is declining in the developed countries while the number of new cases is still increasing in the underdeveloped countries, largely due to the use of contaminated blood for transfusion. The major risk factor for HCV transmission has changed over time from blood transfusion related cases to injecting drug use in the Western world. The relative importance of other risk factors has not changed much over time. These include unsafe sex with multiple partners, occupational and perinatal exposures, nosocomial and iatrogenic infections, unsafe tattooing, piercing and acupuncture, see M.J. Alter, Prevention of spread of hepatitis C, Hepatology 36 (2002) S93-S98.

Hepatitis C is caused by a small RNA virus belonging to the flaviviridae family and has been recently classified as the only member of the genes hepacivirus, see B. Robertson, G. Myers, C. Howard, T. Brettin, J. Bukh, B. Gaschen, et al., Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization, Arch Virol 143 (1998) 2493-2503. The HCV genome is a 9.6 Kb single-stranded RNA which encodes a single polypeptide of about 3000 amino acids, see M. Major, S.M. Feinstone, The molecular biology of

hepatitis C, Hepatology 25 (1997) 1527-1538. This HCV polypeptide is cut post-translationally to generate several structural and non-structural proteins including two envelope glycoproteins (E1 and E2), the nucleopeptide protein (core-C) and several non-structural (from NS2 to NS5) proteins. Some of the viral proteins have been shown to involve in the pathogenesis of the liver disease and also in the development of resistance to interferon therapy. The HCV core proteins, either in its full-length or truncated forms, have been shown to provoke apoptosis of infected cells, see A. Ruggieri, T. Harada, Y. Matsuura, T. Miyamura, Sensitization to Fas-mediated apoptosis by hepatitis C virus core protein, Virology 229 (1997) 68-76, and thus might directly involved in the pathogenesis of liver disease, of cell proliferation and liver cancer development. The core proteins and NS5A have also been reported to interfere with cellular metabolism of lipids and with a direct effect on the development of steatosis, see G. Perlemuter, A. Sabile, P. Letteron, G. Vona, A. Topilco, Y. Chretien, et al., Hepatitis C virus core protein inhibits microsomal triglyceride transfer protein activity and very low density lipoprotein secretion: a model of viral-related steatosis, FASEB J 16 (2002) 185-194, which is a characteristic feature of hepatitis C, see L. Rubbia-Brandt, R. Quadri, K. Abid, E. Giostra, P.J. Male, G. Mentha, et al., Hepatocyte steatosis is a cytopathic effect of hepatitis C virus genotype 3, J Hepatol 33

(2000) 106-115. In addition, NS5A may contain an interferon sensitivity determining region (ISDR) capable of regulating the cellular response to interferon, see N. Enomoto, I. Sakuma, Y. Asahina, M. Kurosaki, T. Murakami, C. Yamamoto, et al., Mutations in the non-structural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection, N Engl J Med. 334 (1996) 77-81. This protein region can bind and inhibit protein kinase R (PKR), whose activity is pivotal for the development of intracellular antiviral state in response to interferon, see S.L. Tan, M.G. Katze, How hepatitis C virus counteracts the interferon response: the jury is still out on NS5A, Virology 284 (2001) 1-12.

There are six major different HCV genotypes and multiple subtypes. Genotypes 1a and 1b are most common in Europe and the United States, followed by genotypes 2 and 3. These four genotypes are also common in the rest of the world. On the other hand, the other genotypes are only common in particular geographic area, such as Egypt in the case of genotype 4, South Africa in the case of genotype 5, and Southeast Asia in the case of genotype 6. The viral genotype is important in terms of the treatment efficacy of antiviral therapy, see T. Poynard, P. Marcellin, S.S. Lee, et al., Randomized trial of interferon alpha 2b plus ribavrin for 48 weeks or for 24 weeks versus interferon alpha 2b plus placebo

for 48 weeks for treatment of chronic infection with hepatitis C virus. Lancet 352 (1998) 1426-32, and J.G. McHutchison, S.C. Gordon, E.R. Schiff, et al., Interferon alpha-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. N Engl J Med. 339 (1998) 1485-92, with better responses associated with genotypes 2 and 3 than with genotype 1. Some of the HCV strains had been reported to have enhanced virulence, but the molecular determinants and mechanisms conferring this property remain elusive, see P. Farci, S.J. Munoz, A. Shimoda, et al., Experimental transmission of hepatitis C virus-associated fulminant hepatitis to a chimpazee. J Infect Dis 179 (1999) 1007-11. Furthermore, genetic variations within a region of NS5A have been deduced to associate with the treatment effectiveness of interferon therapy, as shown in isolates of Japanese subtype 1b, see N. Enomoto, I. Shakuma, Y. Asahina, et al., Mutation in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis virus 1b infection. N Engl J Med. 334 (1996) 77-81. However, this result could not be reproduced in European and American isolates of HCV 1b, see R.T. Chung, A. Monto, J.L. Dienstag, L.M. Kaplan, Mutations in the NS5A region do not predict interferon-responsiveness in American patients infected with genotype 1b hepatitis C virus. J Med Virol 58 (1999) 353-8, and S. Zeuzem, J.H. Lee, W.K. Roth, Mutations in the

non-structural 5A gene of European hepatitis C virus isolates and response to interferon alpha. Hepatology 25 (1997) 740-4.

The natural history of hepatitis C is very heterogeneous, it can either progress towards cirrhosis and its complications, though over a quite long period of time, or remain as benign and one-progressive chronic infection in the majority of the HCV carriers. The severity, progression, and outcome of hepatitis are influenced by several cofactors, see A. Alberti, L. Chemello, L. Benvegnù, Natural history of hepatitis C, J Hepatol 31 (Suppl 1) (1999) 17-24. Retrospective studies conducted in patients with hepatitis C observed for 10-30 years after infection indicate that 17-55% (mean 42%) developed cirrhosis, 1-23% developed HCC and 4-15% died of liver related causes. These figures are quite reduced in most prospective studies where over a follow-up period of 8-16 years after exposure 7-16% of the patients developed cirrhosis (mean 11%), 0.7-1.3% developed HCC and 1.3-3.7% died of liver related causes, see L.B. Seeff, Natural history of chronic hepatitis C, Hepatology 36 (2002) S35-S46. In a series of retrospective-prospective studies lasting 9-45 years, it was found that 0.3-15% of the developed cirrhosis, 0-1.9% HCC, and 0-2.8% died of liver related diseases. These investigations also revealed that many host and environmental factors can influence the course and outcome

of chronic hepatitis C and account for the great heterogeneity of this disease. These differences are very well described by the quite different outcomes and rates of progression to cirrhosis seen when distinct cohort of patients were followed-up for a similar period of time (20-25 years) after infection. In adult patients, mainly males, infected at the age of 45-65 years with a large inoculum through blood transfusion in the pre-serologic era, 15-27% developed cirrhosis, see R.L. Koretz, H. Abbey, E. Coleman, G. Gitnick, NANB post-transfusion hepatitis: looking back on the second decade, Ann Intern Med 119 (1993) 110-115, F. Tremolada, C. Cassin, A. Alberti, C. Drago, A. Tagger, M.L. Ribero, G. Realdi, Long-term follow-up of NANB (type C) post-transfusion hepatitis, J Hepatol 16 (1992) 273-281, and A.M. Di Bisceglie, Z.D. Goodman, K.G. Ishak, J.H. Hoofnagle, J.J. Melpolder, H.J. Alter, Long-term clinical and histopathological follow-up of chronic post-transfusion hepatitis, Hepatology 14 (1991) 969-974, compared to 4% with community-acquired hepatitis C, see A.J. Rodger, S. Roberts, A. Lanigan, S. Bowden, N. Crofts, Assessment of long-term outcomes of community-acquired hepatitis C infection in a cohort with sera stored from 1971-1975, Hepatology 32 (2000) 582-587, 1% of young drug-addicts, see D.L. Thomas, J. Astemborski, R.M. Rai, F.A. Anania, M. Schaeffer, N. Galai, et al., The natural history of hepatitis C virus infection: host, viral and

environmental factors, J Am Med Assoc 284 (2000) 450-456, 0.4-2% of young women contaminated by anti-D Ig preparations, see E. Kenny-Walsh, for the Irish Hepatology Research Group, Clinical outcome after hepatitis C infection from contaminated anti-D immune globulin, N Engl J Med 340 (1999) 1228-1233, and 0.3% of children with hepatitis C, see M. Wiese, F. Berr, M. Lafrenz, H. Porst, V. Olsen, Low frequency of cirrhosis in a hepatitis C (genotype 1b) single-source outbreak in Germany: a 20-year multicenter study, Hepatology 32 (2000) 91-96. These findings indicate that size and source of infection, age and gender are important variables affecting the course and outcome of chronic hepatitis C.

The treatment of patients with chronic HCV infection is based largely on consensus guidelines, see National Institutes of Health Consensus Development Conference Panel statement: management of hepatitis C. Hepatology 26 (1997) Suppl 1:2S-10S, and EASL International Consensus Conference on Hepatitis C: Paris, 26-28, February 1999, consensus statement. J Hepatol 30 (1999) 956-961. The 1999 recommendations, see EASL International Consensus Conference on Hepatitis C: Paris, 26-28, February 1999, consensus statement. J Hepatol 30 (1999) 956-961, suggest that naïve patients with the above-described

indications and without contraindications to treatment with interferon or ribavirin should receive combination therapy. Treatment consists of 3 million U of interferon-α administered subsetaneously three times a week and 1200 mg of ribavirin orally per day for patients with weight greater than 75 kg and 1000 mg of ribavirin for those less than 75 kg. Usually, ribavirin is taken in divided doses, given in the morning and evening, and interferon is given before bedtime.

The efficiency of these therapies is usually and conventionally determined by measuring a biochemical response (normalization of alanine aminotransferase levels), but recently the introduction of assays for the detection of HCV RNA have allowed the assessment of virologic response (as defined by a negative result on a qualitative PCR assay for HCV RNA) as a criteria for successful therapy as well. Since responses to therapy may not be maintained after treatment is stopped, the success of clinical trials has been evaluated in terms of the response at the end of therapy (end-of-treatment response) and six months after the cessation of treatment (sustained treatment response). Patients with a sustained virologic response have a high probability of having a durable biochemical, virologic, and histologic response, see O. Reichard, H. Glaumann, A. Fryden, G. Norkrans, R. Wejstal, O. Weiland, Long-term follow-up of chronic hepatitis C patients

with sustained virological response to alpha-interferon. J Hepatol 30 (1999) 783-787.

The rate of end-of-treatment response of HCV patients to interferon monotherapy was as high as 40 percent, but the rate of sustained response is less than half of this, see T. Poynard, P. Marcellin, S.S. Lee, et al. Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. Lancet 352 (1998) 1426-1432, and J.G. McHutchison, S.C. Gordon, E.R. Schiff, et al. Interferon alfa-2b alone or in combination with ribavirin as initital treatment for chronic hepatitis C. N Engl J Med 339 (1998) 1485-1492. This is especially true in persons infected with HCV genotype 1a or 1b, the most prevalent genotypes in the United States and western Europe. Two large, prospective trials, see T. Poynard, P. Marcellin, S.S. Lee, et al., Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. Lancet 352 (1998) 1426-1432, and J.G. McHutchison, S.C. Gordon, E.R. Schiff, et al., Interferon alfa-2b alone or in combination with ribavirin as initital treatment for chronic hepatitis C. N Engl J Med 339 (1998) 1485-1492, demonstrated that the combination of

interferon-α and ribavirin significantly elevates the percentage of naive patients who have a sustained virologic response, from 16% to 40%. Also, both studies showed that the treatment regimens with optimal clinical outcome were associated with the viral genotype and pre-treatment viral load. For patients infected with HCV genotype 2 or 3 and in those with low viral loads before treatment, the response was maximal after 24 weeks of the treatment, whereas patients infected with genotype 1 and those with a high viral load before treatment required a course of 48 weeks for an optimal outcome. This finding led to the recommendation that the duration of treatment should be based on the HCV genotype and the pretreatment viral load, see EASL International Consensus Conference on Hepatitis C: Paris, 26-28, February 1999, consensus statement. J Hepatol 30 (1999) 956-961. However, since tests for the quantification of HCV RNA are still not standardized, and since the viral load naturally fluctuates over time, the viral load is currently not routinely used for determining the treatment regimen.

Therapy of chronic hepatitis C has greatly improved in recent years with the use of interferon-α and ribavirin combined therapy. The therapy has been further improved more recently with the use of pegylated interferons (PEG-IFNs), again combined with ribavirin. The recent NIH Consensus

Conference of the Management of Hepatitis C has concluded that on the basis of available data the highest response rates to antiviral therapy for the treatment of chronic hepatitis C have been achieved using the combination of PEG-IFNs and ribavirin, at least for patients infected with HCV-1 and such regimen has been therefore proposed as the new standard of therapy for chronic hepatitis, see C A. Alberti and L. Benvegnu, Management of hepatitis C, J. Hepatology 38 (2003) S104-S118.

Factors influencing the rate of sustained virological response include viral and host factors, as well as the pathogenesis state of the liver. The viral factors include viral genotype (types 1a, 2 and 3 are favorable to response), level of viraemia (less than 2 million copies/ml is favorable), and level of viral heterogeneity (degree of variability in E2/NS1 region of HCV correlates with response to IFN). The favorable host factors include younger age (less than 40) as well as female sex. The beneficial pathogenic factors are lower ALT and AST levels before treatment, absence of cirrhosis and low fibrotic histological scores, and lower hepatic iron content, see A. Alberti and L. Benvegnu, Management of hepatitis C, J. Hepatology 38 (2003) S104-S118, and G.M. Lauer and B.D. Walker, Hepatitis C virus infection, N Engl J Med 345 (2001) 41-52. In addition, due to the great extent of inter-individual

variations in response to the treatment, it has been speculated that host genetic factors may also play an important role.

CD81 is a membrane bound protein composed of four transmembrane and two extracellular domains with a molecular weight of 26-kDa, see S. Levy, S.C. Todd, and H.T. Maecker, CD81 (TAPA-1): A molecule involved in signal transduction and cell adhesion in the immune system. Annu. Rev. Immunol. 16 (1998) 89-109. It is a member of the superfamily of proteins known as tetraspanins, see H.T. Maecker, S.C. Todd, and S. Levy, The tetraspanin superfamily: molecular facilitators. FASEB J. 11 (1997) 428-442. Most tetraspanins were originally identified as leukocyte antigens; however, it is now becoming evident that generally tetraspanins, particularly CD81, is expressed in many different cell types and involved in a variety of cellular functions including cell adhesion and migration, alteration of cell morphology, and activation state of a cell, see I. Tachibana and M. E. Hemler, Role of transmembrane 4 superfamily (TM4SF) proteins CD9 and CD81 in muscle cell fusion and myotube maintenance, J. Cell Biol. 146 (1999) 893-904. In the immune system, on the B cell CD81 forms a complex with CD21, CD19, and Leu13. Formation of this complex decreases the threshold for B cell activation through the B cell receptor by bridging Ag specific recognition and

CD21-mediated complement recognition, see D.T. Fearon and R.H. Carter, The CD19/CR2/TAPA-1 complex of B lymphocytes: linking natural to acquired immunity. Annu. Rev. Immunol. 13 (1995) 127-149. On T cells CD81 associates with CD4 and CD8 and provides a costimulatory signal with CD3, see T. Imai, M. Kakizaki, M. Nishimura, and O. Yoshie, Molecular analyses if the association of CD4 with two members of the transmembrane 4 superfamily, CD81 and CD82. J. Immunol. 15 (1995) 1229-1239. It is also shown that expression of CD81 by T cells greatly enhances cognate T-B cell interactions and greatly amplifies Th2 polarized intracellular activation pathways, see J. Deng, R.H. Dekruyff, G.J. Freeman, D.T. Umetsu, and S. Levy, Critical role of CD81 in cognate T-B cell interactions leading to Th2 responses, Intl. Immunol. 14 (2002) 513-523.

Recently, it was shown that HCV particles bind CD81 and this binding is mediated by the interaction of the second extracellular loop of CD81 with HCV envelope 2 glycoprotien in vitro, see P. Pileri, Y. Uematsu, S. Campagnoli, G. Galli, F. Falugi, R. Petracca, A.J. Weiner, M. Houghton, D. Rosa, G. Grandi, and S. Abrignani, Binding of hepatitis C virus to CD81. Science 282 (1998) 938-941. Since the interaction between CD81 and E2 is sufficient for binding of whole HCV particle, it was postulated that CD81 may act as a

receptor for the attachment and entrance of HCV into the cell, see M. Flint, C. Maidens, L.D. Loomis-Price, C. Shotton, J. Dubuisson, P. Monk, A. Higginbottom, S Levy, and J. A. McKeating, Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor CD81. J. Virol. 73 (1999) 6235-6244. However, until now there is no conclusive evidence to demonstrate the involvement of CD81 in the cellular uptake of HCV virions. More recently, it was showed that the protein level of membrane bound CD81 in isolated human peripheral blood cells and hepatocytes was significantly down-regulated by the treatment of interferon-α alone or combined with ribavirin, see B. Kronenberger, B. Ruster, R. Liez, S. Weber, A. Piier, J.H. Lee, W.K. Roth, and S. Zeuzem, Interferon alfa down-regulates CD81 in patients with chronic hepatitis C. Hepatology 33 (2001) 1518-1526. Also, levels of total CD81 protein of the PBLs of HCV-infected patients are significantly higher than those of the healthy subjects. Furthermore, cell surface-associated CD81 protein was lower 4 weeks after initiation of therapy in patients with an initial virologic response compared with initial virologic non-responders. Therefore, it is concluded that interferon-α and ribavirin regulate the expression of CD81 in vitro and in vivo. CD81 expression correlates with initial virologic response in patients with HCV infection. However, the detailed

regulatory mechanism of CD 81 expression by interferon-α and ribavirin remains unclear. Moreover, the underlying reason that differentiates patients with different level of CD81 expression in response to therapy requires further studies at the genomic level, particularly for the sustained response of the therapy.

SUMMARY OF THE INVENTION

Accordingly, the present invention provides isolated polynucleotides that encode sequences for CD81 gene, which is shown to be associated with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy. The polynucleotides include polymorphisms associated with responsiveness of interferon-α and ribavirin and are useful as the probes in screening for HCV infected patients suitable for interferon-α and ribavirin combined therapy. The present invention also provides linkage disequilibrium structure of CD81, haplotype information and its use for prediction of potential responders. The present invention further provides methods for detecting polymorphisms in CD81 gene and its surrounding regions, and methods of detecting a propensity to response to the therapy of interferon-α and ribavirin, using the isolated polynucleotides of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other objects, features and advantages of the present invention will become apparent to those skilled in the art upon consideration of the following description of the preferred embodiments of the present invention taken in conjunction with the accompanying drawings, in which:

Fig. 1 shows Table 1 of the primers for SNP genotyping of CD81 SNPs with FP-TDI method;

Fig. 2 shows the sequence traces of the responder and non-responder at and around rs800136;

Fig. 3 shows the sequence traces of the responder and non-responder at and around rs800137;

Fig. 4 shows the sequence traces of the responder and non-responder at and around rs800334;

Fig. 5 shows the sequence traces of the responder and non-responder at and around pos1989603;

Fig. 6 shows the sequence traces of the responder and non-responder at and around rs8002522012 and rs8002522013;

Fig. 7 shows the sequence traces of the responder and non-responder at and around rs800335;

Fig. 8 shows Table 2 of the distribution of CD81 SNPs in Chinese HCV infected patients;

Fig. 9 shows Table 3 of the association of SNP markers of CD81 gene and its flanking regions with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy;

Fig. 10 shows Table 4 of the CD81-SNPrs800136 alleles and associated treatment response status;

Fig. 11 shows Table 5 of the CD81-SNPrs800137 alleles and associated treatment response status;

Fig. 12 shows Tables 6 and 7 of the CD81-SNPrs800334 alleles and genotypes and associated treatment response status;

Fig. 13 shows Tables 8 and 9 of the CD81-SNPpos1989603 alleles and genotypes and associated treatment response status;

Fig. 14 shows Tables 10 and 11 of the CD81-SNPrs2522012 alleles and genotypes and associated treatment response status;

Fig. 15 shows Tables 12 and 13 of the CD81-SNPrs2522013 alleles and genotypes and associated treatment response status;

Fig. 16 shows Tables 14 and 15 of the CD81-SNPrs800335 alleles and genotypes and associated treatment response status;

Fig. 17 shows the haplotype blocks that encompass the 5', 3', and the intragenic region of CD 81 gene, and SNPs showing to be associated with response to interferon-a and ribavirin combined therapy;

Fig. 18 shows Table 16 of the CD81-SNPrs800136 and rs800137 haplotype and associated treatment response status; and

Fig. 19 shows Table 17 of the CD81-SNPrs800334, pos1989603, rs2522012, rs2522013, rs800335 haplotypes and associated treatment response status.

DETAILED DESCRIPTION OF THE INVENTION

It is directed to the investigation of the correlation of clinical responsiveness of HCV infected patients to interferon-α and ribavirin combined therapy with genetic polymorphisms in and surrounding the CD81 gene. The results indicate that the treatment responsiveness is associated with the host genotype on CD81 gene. The distribution of the allele and genotype of several SNPs upstream the CD81 gene is significantly different between the responder and non-responder groups. Further analysis of the linkage disequilibrium structure of the CD81 gene demonstrates that the significant SNPs are clustered in two distinct LD blocks. Moreover, distribution of haplotypes in these two blocks is also significantly different between the responder and non-responder groups. All of these results indicate that CD81 may directly involve in the treatment response pathway, and its genetic variations play an important role in determining the therapeutic outcome.

For reference, Appendix shows the nucleotide sequences of CD81 gene and its 5'-flanking region extended 5K upstream of exon 1 and the 3'-flanking region extended 1 kb downstream of the poly-A tail.

Methods and materials

<Study subjects>

Ninety two Chinese Han patients with chronic HCV infection were recruited at National Taiwan University Hospital. All patients' blood samples were Anti-HCV(+) and HCV RNA (+). Patients were excluded from receiving interferon therapy if they had any of the following criteria: neutrophil count < 1,500 cells/mm^3, Hgb <12g/dL in women or 13 g/dL in men, or platelet count < 90,000 cells/mm^3, history of poorly controlled thyroid disease, and serum creatinine level > 1.5 times the upper limit of normal at screening.

All patients received a combination therapy of ribavirin at a dose of 1200mg per day and interferon-alpha (2a or 2b) at a dosage of 3-5 MU 3 times per week for 6 months, and were subsequently followed for treatment response via clinical, biochemical, and serologic markers for more than one

year.

The definition of sustained responders to IFN and ribavirin combination treatment for chronic hepatitis C disease included patients with HCV RNA (+) to HCV RNA (-) conversion 6 months after treatment period. Patients with concurrent hepatitis B or D infection are excluded. Informed consent was obtained in writing from each patient. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki as reflected by approval from the institutional review committee.

<Isolation of genomic DNA>

Genomic DNA is isolated from blood sample using QIAamp DNA Blood kit according to the manufacture's instructions. The isolated genomic DNA is quality check by agarose gel electrophoresis analysis, quantity determined spectrophotometrically, and stored at -80 °C until use.

<Analysis of single nucleotide polymorhpism>

SNP analysis was performed by either direct DNA sequencing or template-directed dye-terminator incorporation with fluorescence-polarization detection (FP-TDI) developed

originally by Chen et al. (1999), see X. Chen, L. Levine, and P.-Y. Kwok, Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res. 9 (1999) 492-8.

SNP analysis by direct DNA sequencing

Multiplex amplification of CD81 fragment from genomic DNA

The fragment of CD81 gene are amplified by a two-step PCR reaction. The initial amplification step is a multiplex PCR reaction containing 12 different pairs of PCR primers. The reaction mixture consists of 50 ng genomic DNA, 0.1 μM each of 12 pairs of primer, 0.25 mM dNTP mixture, 100 mM KCl, 20 mM Tris-HCl pH 8.3, 0.2% Triton X-100, and 5 mM $MgCl_2$, 10U of Vio*Taq* DNA polymerase (VIOGENE) and 0.05U of *pfu* DNA polymerase (Stratagene) in a total volume of 100 μL reaction. The reaction is performed by a touchdown program with an initial denaturing at 94°C for 4 min, 10 cycles of melting at 94°C for 40 sec, annealing at 72°C with 1°C decrement per cycle for 40 sec, and extending at 72°C for 1 min 30 sec; for the subsequent 25 cycles, the annealing temperature is 62°C with the same conditions for denaturing and extending procedures, and one cycle of final extension at 72°C for 10 min. Amplification is carried out using 2700 PCR machines (ABI) and the amplified products are purified by

membrane ultra-filtration with MultiScreen PCR plate (Millipore) according to the manufacture's instruction. In the next step, specific 791 bp of CD81 product is amplified using the purified multiplex product as template is amplified from the simultaneously amplified products in a 78 to 72 touchdown program as described previously in 1 fold of PCR buffer. U.S. Pat. Application No. 10/446,940 is also attached hereto for more detail to perform a two-step PCR reaction with touchdown programs.

Sequencing PCR products using dye-labeled terminators

Each amplified and purified reaction product is divided into 96 sequencing reactions. The sequencing reaction is performed in a PCR machine with each reaction mixture consisting of PCR product, Big Dye Terminator Ready-Reaction-Premix and 10 pmol of sequencing primer. Reaction is subjected to 28 cycles at 94 °C for 30 sec, 48 °C for 30 sec, and 58 °C for 2 min. Reaction product is purified by ethanol precipitation, re-suspended in ultra-pure water, and loaded on an ABI 3700 capillary sequencer.

Sequence analysis and allele calling

DNA sequence data and track are collected by ABI

DNA Sequence Analyzer. Base calling is performed with Phred and sequencing assembled by Phrad. For ensuring the data quality only bases with Q value greater than 20 are analyzed further. PolyPhred is used for the initial identification of potential SNP sites. SNP sites identified are verified manually by three experienced researchers independently, and those scored by at least two researchers are deemed as true SNPs. Similarly, genotype of each tested individual is initially determined by the computer software and confirmed manually as for the identification of SNP site.

The work flow for SNP genotyping by TDI-FP is as follow: (1) primer design, (2) PCR amplification, (3) excess PCR primer and dNTP degradation, (4) single-base extension, and (5) FP measurement.

(1) Primer design

Four primers are designed for each SNP site, two for PCR amplification of the DNA fragment containing the SNP site and two for TDI reaction. Primer 3 is employed to design the PCR primers. The PCR primers are designed to have a melting temperature between 54 °C to 56 °C. The TDI primers are designed by a program, developed originally by Vieux et al., see E. F. Vieux, P.-Y. Kwok, R. D. Miller, Primer

design for PCR and sequencing in high-throughput analysis of SNPs. Biotechniques. (2002) Suppl: 28-30, 32., and modified in house by our bioinformatics group, to have melting temperature between 50 °C to 55 °C and lengths between 20 to 30 bases (about 10,000 Da). Primers used for genotyping of SNPs of CD81 in this study are listed in Table 1 of Fig. 1.

(2) PCR amplification

Human genomic DNA (5 ng) is amplified in a 10 μl reaction mixtures containing 1x PCR buffer (Applied Biosystems), 2.5 mM $MgCl_2$, 50 μM dNTP, 0.1 μM each of forward and reverse primers, and 0.2 U AmpliTaq Gold DNA polymerase. The reaction mixture is held at 95 °C for 10 minutes for activation of AmpliTaq Gold DNA polymerase, followed by 35 amplification cycles. Each cycle consists of denaturation at 95 °C for 10 seconds, primer annealing at 55 °C for 20 seconds, and primer extension at 72 °C for 30 seconds. At the end of the thermal cycling, the reaction is held at 72 °C for 10 minutes for final primer extension and then incubated at 4 °C until use.

(3) Excess PCR primer and dNTP degradation

After completion of the PCR reaction, a 2-μl PCR

Clean-Up reagent is added to 5 μl of the PCR reaction mixture. The PCR Clean-Up reagent contains 1 U shrimp alkaline phosphatase, and 1 U *E. coli* exonuclease I in shrimp alkaline phosphatase buffer (25 mM Tris-HCl, pH 8.5 and 2.5 mM MgCl2). The degradation reaction is performed at 37 °C for one hour followed by an incubation at 80 °C for 15 minutes for inactivation of the enzymatic activity.

(4) Single-base extension

To the Exo-SAP enzymatic treated PCR product (7 μl), 13 μl of the TDI reaction cocktail (AcycloPrime-FP Mix, Perkin Elmer) is added. The cocktail consists of components to give the reaction at the final concentrations of 1x reaction buffer (Perkin Elmer), 0.25 M SNP primer, 1 μl Acyclo Terminator Mix (Perkin Elmer), and 0.25 U AcycloPol DNA polymerase (Perkin Elmer). The reaction is incubated at 95 °C for 2 minutes followed by 25 to 50 cycles of 95 °C for 15 seconds and 55 °C for 30 seconds. At the end of the reaction, the samples are held at 4 °C.

(5) FP measurement

FP measurement is conducted on a Perkin Elmer fluorescence reader (Victor^2). FP value is defined by the

formula

$$p = (Ivv - Ivh) / (Ivv + Ivh), \qquad [Eq.1]$$

where Ivv is the emission intensity measured when the excitation and emission polarized filters are parallel and Ivh is the emission intensity measured when these filters are perpendicular. The measurements are mathematically transformed to the expression of millipolarization

$$mP = 1000 \times (Ivv - (G \times Ivh)) / (Ivv + (G \times Ivh)), \qquad [Eq.2]$$

where G is a gating factor near the value of one used to adjust for slight differences in the two optical paths used. G is estimated for each of the four dyes using an assumed mP value for the dye of 30 mP and measured values for Ivv and Ivh.

Determination of genotype

The identification of the genotype of each sample is assisted by the allele calling program SNP Scorer provided by the manufacturer. In brief, the mP values of the two dyes, representing one of the two possible alleles, are plotted on an abscissa. A data clustering algorithm is used for

classification of the data points into four separate groups. The genotype of each sample is assigned dependent on the group it belongs to.

Figs. 2-7 show the sequence traces of the responder and non-responder at and around various CD81-SNP alleles, respectively. Fig. 2 shows the sequence traces of the responder and non-responder at and around rs800136. Fig. 3 shows the sequence traces of the responder and non-responder at and around rs800137. Fig. 4 shows the sequence traces of the responder and non-responder at and around rs800334. Fig. 5 shows the sequence traces of the responder and non-responder at and around pos1989603. Fig. 6 shows the sequence traces of the responder and non-responder at and around rs8002522012 and rs8002522013. Fig. 7 shows the sequence traces of the responder and non-responder at and around rs800335.

Estimation of haplotypes and frequencies

Haplotypes and their frequencies are estimated on the basis of unphased genotype data, by the computer program PHASE. Described in the work of Stephens et al. (2001), PHASE uses a Bayesian algorithm that regards the unknown haplotypes as unobserved random quantities and

aims to evaluate their frequencies and conditional distribution of multilocus haplotypes in diploid populations.

Statistical analysis of linkage disequilibrium

Initial LD analysis is computed by performance of pair-wise comparisons for all SNP loci. Significance levels are determined by the χ^2 test. SNPs having a minor-allele frequency less than 0.05 are excluded from LD analysis for the corresponding 2×2 table. Several widely used LD statistics of linkage disequilibrium are also computed: D, D' and r^2. For two loci L_1 and L_2, each with two alleles 1 and 2, let p_i be the frequency of allele 1 and $q_i=1-p_i$ be the frequency of allele 2, at locus i (i=1, 2). Let p_{11} be the frequency of the 11 haplotype and in general let p_{jk} be the frequency of the jk haplotype. The coefficient of disequilibrium, D, is the difference between the observed haplotype frequency and the frequency expected under statistical independence: $D=p_{11}-p_1 p_2$. The labeling of the alleles may affect the sign of D, but not its absolute value. The relative disequilibrium D' ranged from 1 to -1, is obtained by dividing D by its possible maximal possible (absolute) value ($D'=D/|D|_{max}$, where $|D|_{max}=\max(p_1p_2, q_1q_2)$ if $D<0$ and $|D|_{max}=\min(q_1p_2, p_1q_2)$ if $D>0$) and is independent of allele frequencies. The correlation coefficient is $r^2 = D^2/p_1p_2q_1q_2$. The study reports D', r^2, χ^2, and *P*-values corresponding to χ^2.

Statistical analysis of association study

Genotype frequencies of each SNP and allele frequencies between R and NR will be compared by using the χ^2 test or Fisher exact test. R as a reference group, the odd ratio is performed to test which allele or genotype is risking factor. According to odd ratio, alleles and genotypes associated with responsiveness or non-responsiveness are defined. To evaluate the combined genotypes of multiple SNPs in the same genes or pathway, the χ^2 test or Fisher exact test is also performed. All statistical tests are 2-tailed, and *P*-values less than 0.05 are considered statistically significant. The analyses are performed using the SAS statistical package version 8.

Results

(1) Selected SNPs of CD81 gene

Over the selected 70 Kb chromosomal region, including CD81 gene and extending 30 Kb each to the 5' and 3' flanking regions, eighteen SNPs and one insertion are identified among HCV infected patients of Chinese population. The positions, sequences, and allele frequencies are

summarized in Table 2 of Fig. 8. Among theses polymorphic sites, 16 SNPs have minor allele frequency greater than 10% and considered to be informative markers for the association analysis of genetic polymorphism with treatment response.

(2) Association of SNP markers of CD81 gene and its flanking regions with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Also shown in Table 3 of Fig. 9, among the 19 SNPs, 7 show allelic association with the treatment response status. In five of these seven SNPs the genotype distribution between responder and non-responder are statistically different. The results of statistic analysis are summarized in Tables 4 to 14 in the below.

(3) Association of CD81-SNPrs800136 allele with treatment responsiveness of HCV patients to interferon-□ and ribavirin combined therapy

Distribution of CD81-SNPrs800136 alleles and the status of treatment response are summarized in Table 4 of Fig. 10. The CD81-SNPrs800136 T allele is found to be associated with responding status and thus a favorable allele for drug treatment.

(4) Association of CD81-SNPrs800137 allele with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPrs800137 alleles and the status of treatment response are summarized in Table 5 of Fig. 11. The CD81-SNPrs800137 T allele is found to be associated with responding status and thus a favorable allele for drug treatment.

(5) Association of CD81-SNPrs800334 with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPrs800334 alleles, genotypes, and the status of treatment response are summarized in Tables 6 and 7 of Fig. 12. The CD81-SNPrs800334 G allele and the GG genotype are found to be associated with responding status and thus a favorable allele for drug treatment.

(6) Association of CD81-SNPpos1989603 with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPpos1989603 alleles, genotypes, and the status of treatment response are summarized in Tables 8 and 9 of Fig. 13. The CD81-SNPpos1989603 A allele and the AA genotype are found to be associated with responding status and thus a favorable allele for drug treatment.

(7) Association of CD81-SNPrs2522012 with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPrs2522012 alleles, genotypes, and the status of treatment response are summarized in Tables 10 and 11 of Fig. 14. The CD81-SNPrs2522012 T allele and the TT genotype are found to be associated with responding status and thus a favorable allele for drug treatment.

(8) Association of CD81-SNPrs2522013 with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPrs2522013 alleles, genotypes, and the status of treatment response are

summarized in Tables 12 and 13 of Fig. 15. The CD81-SNPrs2522013 A allele and the AA genotype are found to be associated with responding status and thus a favorable allele for drug treatment.

(9) Association of CD81-SNPrs800335 with treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Distribution of CD81-SNPrs800335 alleles, genotypes, and the status of treatment response are summarized in Tables 14 and 15 of Fig. 16. The CD81-SNPrs800335 T allele and the TT genotype are found to be associated with responding status and thus a favorable allele for drug treatment.

(10) Linkage disequilibrium (LD) structure of CD81 gene

The LD structure of CD81 gene and its flanking regions extended 30 Kb each to the 5' and 3' ends are constructed by genotypes of the selected 16 informative SNPs. In this 70 Kb chromosomal fragment of DNA, 9 haplotype blocks are identified as depicted in Fig. 17.

(11) Association of haplotypes in haplotype block 4 with

treatment responsiveness of HCV patients to interferon-α and ribavirin combined therapy

Two SNPs are located in haplotype block 4, and the distribution of the haplotype and the status of treatment response are summarized in Table 16 of Fig. 18. The GG haplotype is found to be associated with responding status and thus a favorable haplotype for drug treatment.

(12) Association of haplotypes in haplotype block 7 with treatment responsiveness of HCV patients to interferon-□ and ribavirin combined therapy

Haplotypes of the seventh haplotype block are constructed by the five informative SNP markers. Distribution of the haplotype and the status of treatment response are summarized in Table 17 of Fig. 19. The TGGCC haplotype is found to be associated with responding status and thus a favorable haplotype for drug treatment. In contrast, the GATAT haplotype is found to be associated with non-responsiveness and a non-favorable factor for drug treatment.

While the present invention has been described in conjunction with preferred embodiments thereof, it is evident

that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and scope thereof as set forth in the appended claims.